

Human inhibin genes

Genomic characterisation and sequencing

Andrew G. Stewart, Helen M. Milborrow, Jennifer M. Ring, Carol E. Crowther and Robert G. Forage

Biotechnology Australia Pty Ltd, PO Box 20, East Roseville, New South Wales 2069, Australia

Received 23 June 1986; revised version received 11 August 1986

Inhibin is a gonadal hormone involved in the non-steroidal regulation of follicle stimulating hormone (FSH) secretion. Using the cDNAs coding for bovine inhibin A and B subunits we have identified inhibin genes within the human genome using Southern blot hybridisation techniques. The genes are likely to be present as single copies. Cloning and sequencing inhibin genes obtained from lambda libraries of human genomic DNA provide structural and sequence data on the human A and B genes. Comparison of the known inhibin gene sequences showed, in particular, that the B subunits have identical sequences in man, pigs and cattle thus demonstrating a remarkable evolutionary conservation in these genes.

Inhibin DNA hybridization Gene cloning Reproductive hormone

1. INTRODUCTION

The controversy surrounding the existence of inhibin, a hormone postulated to be produced in the gonads and involved in the feedback control of FSH secretion [1,2], has been clarified with the purification of inhibin from two animal species [3,4]. Bovine follicular fluid yielded a protein with a molecular mass of 58 kDa composed of two disulphide-linked subunits termed A and B of 43 and 15 kDa, respectively [3]. Subsequently a smaller active form (31–32 kDa) was isolated from both porcine and bovine follicular fluids with subunits A_c and B of 20 and 15 kDa, respectively [4–8]. These observations were confirmed by more detailed information on the structures of inhibin, its subunits and precursors obtained by the recent cloning of the cDNAs for the subunits of bovine, porcine, human and bovine inhibins [9–11].

This study describes the cloning and characterisation of genomic sequences of human inhibin A and B genes and examines the

homologies between three mammalian inhibin species.

2. MATERIALS AND METHODS

2.1. Southern blot analysis

Human genomic DNA from a human macrophage cell line U937 (ATCC CRL 1539) was digested with various restriction enzymes, subjected to electrophoresis on agarose gels and transferred to nitrocellulose membranes [12]. These were prehybridised by incubating in 50% formamide, 5 × SSPE, 5 × Denhardt's solution and 25 µg/ml denatured herring sperm DNA for 4 h at 42°C, then probed with [³²P]cDNA coding for substantial parts of the 43 and 15 kDa bovine inhibin subunits. For the A subunit a 725 base pair (bp) fragment of the bovine inhibin cDNA A chain gene was used (bases 242–967 [11]). For the B subunit a 358 bp fragment of the bovine cDNA B chain gene was used (bases 1–358 [11]). Labelling was performed by the oligonucleotide method [13]

and hybridisation was as for the prehybridisation but in the added presence of 10% dextran sulphate and only $1 \times$ Denhardt's solution. Filters were washed in $0.2 \times$ SSC and 0.1% SDS at 50°C [12] and radioactive bands detected by autoradiography.

2.2. Isolation of human inhibin genes

Two λ bacteriophage libraries of human genomic DNA were used. One, in λ L47 [14] was used for the isolation of the A gene, the other, in the λ EMBL3 vector [15] was used for the preparation of the B gene. Plaques were transferred to nitrocellulose filters and probed with bovine inhibin cDNA fragments as described for the Southern Blot analysis. Positive plaques were purified, the DNA extracted and the insert mapped for restriction sites and partially sequenced by the dideoxynucleotide chain-termination method [16].

3. RESULTS

3.1. Inhibin sequences in the human genome

Sequences similar to or homologous with the genes coding for the large (43 kDa) or the small (15 kDa) subunits of bovine inhibin were identified in human genomic DNA by standard Southern Blot hybridisation techniques [12]. A number of fragments hybridised to cDNA coding for the large or small subunits of bovine inhibin as shown in table 1. The conditions of hybridisation and washing were such that the hybridising human DNA must be at least 75% homologous overall

with the bovine cDNA. Digestion of the DNA with most of the enzymes (except *Pst*I) resulted in the detection of a single fragment which hybridised to either the large or the small subunit probes. Human DNA, therefore, contains sequences which are at least 75% homologous with bovine inhibin genes and may represent the human inhibin genes. These genes are present in low copy number, probably unique and may not be adjacent since the probes for the large and small subunits did not hybridise to similar sized fragments.

3.2. Isolation of inhibin genes from human genomic DNA

Phages hybridising to the bovine inhibin cDNAs were isolated from libraries of human genomic DNA in bacteriophage λ . One clone, λ BTA1406, contained an insert of human DNA of 11.5 kb. Digestion with *Pst*I yielded three fragments of about 2550, 1160 and 480 bp, respectively, which hybridised to the bovine inhibin A chain gene. These are the same fragments observed by Southern blot analysis of *Pst*I digested genomic DNA (table 1). The nucleotide sequence of part of this insert was determined and identified it as the gene for the preproA subunit of human inhibin by homology with the bovine sequence (fig.1). In this gene there is one intron inserted in the codon for amino acid 29 which is surrounded by typical intron/exon boundaries [17]. This intron is in a similar position to one found in an unspliced bovine inhibin A chain mRNA [11]. The complete coding sequence of the human preproinhibin A chain gene is shown in fig.1. This data extends the previously published sequence which differs slightly in the N-terminal region [10]. Also shown is a comparison between the amino acid sequences of human, bovine and porcine preproinhibins which are about 84% homologous. The overall structure of the A gene is shown in fig.2.

Another clone, λ BTA1407, which hybridised to the bovine B subunit gene, contained an insert of human DNA of about 18 kb. The nucleotide sequence of part of this insert coded for the B chain of human inhibin by its homology with bovine and porcine inhibin genes (fig.3). There is a 9.7 kb intron 542 bp upstream from base position 1 as shown in fig.2. The B proteins encoded by all three genes are identical.

Table 1

Hybridisation of bovine inhibin cDNA to human genomic DNA

Probe	Enzymes	Fragment sizes (bp)
Bovine large subunit cDNA	<i>Pst</i> I	2450, 1200, 480
	<i>Pvu</i> II	700
	<i>Bam</i> HI	18000
	<i>Eco</i> RI	15000
	<i>Hind</i> III	20000
Bovine small subunit cDNA	<i>Pst</i> I	650
	<i>Eco</i> RI	9000
	<i>Hind</i> III	11000

Human	1	gtgagct	ATG	GTG	CTG	CAC	CTA	CTG	CTG	TTC	TTG	CTG	CTG	ACC	CCA	CAG	GGT	GGG	CAC	AGC	TGC	CAG	GGG	CTG	GAG	CTG	GCC	CGG	GAA		
Human			Met	Val	Leu	His	Leu	Leu	Phe	Leu	Leu	Leu	Leu	Thr	Pro	Gln	Gly	Gly	His	Ser	Cys	Gln	Gly	Leu	Glu	Leu	Ala	Arg	Glu		
Bovine				Trp		Gln	---			Leu				Ala						Gly	His						Asp				
Porcine				Trp	Pro	Gln	---			Leu				Ala		Arg	Ser			Gly			Pro				Asp				
Human	89	CTT GTT	CTG	GCC	AAG	GTG	AGG	GCC	CTG	TTC	TTG	GAT	GCC	TTG	GGG	CCC	CCC	GCG	GTG	ACC	AGG	GAA	GGT	GGG	GAC	CCT	GGA	GTC	AGG		
Human		Leu Val	Leu	Ala	Lys	Val	Arg	Ala	Leu	Phe	Leu	Asp	Ala	Leu	Gly	Pro	Pro	Ala	Val	Thr	Arg	Glu	Gly	Gly	Asp	Pro	Gly	Val	Arg		
Bovine																		Pro		Arg	Gly										
Porcine																															
Human	176	CGG	CTG	CCC	CGA	AGA	CAT	GCC	CTG	GGG	GGC	TTC	ACA	CAC	AGG	GGC	TCT	GAG	CCC	GAG	GAA	GAG	GAG	GAT	GTC	TCC	CAA	GCC	ATC	CTT	
Human		Arg	Leu	Pro	Arg	Arg	His	Ala	Leu	Gly	Gly	Phe	Met	Arg	Arg	Gly	Ser	Glu	Pro	Glu	Glu	Glu	Glu	Asp	Val	Ser	Gln	Ala	Ile	Leu	
Bovine				His					Val				Val								Asp	Gln	---								
Porcine									Val				Met										---								
Human	263	TTC	CCA	GCC	ACA	GgtaacggcagAT	30	AGC	TGT	GAG	GAC	AAG	TCA	GCT	GCC	AGA	GGG	CTG	GCC	CAG	GAG	GCT	GAG	GCT	GAG	GAG	GCC			
Human		Phe	Pro	Ala	Thr	A	intron	sp	Ala	Ser	Cys	Glu	Asp	Lys	Ser	Ala	Arg	Gly	Leu	Ala	Gln	Gly	Leu	Ala	Gln	Glu	Ala	Glu	Gly		
Bovine					Ala			Gly				Arg							---	---	---										
Porcine																															
Human	345	CTC	TTC	AGA	TAC	ATG	TTC	CGG	CCA	TCC	CAG	CAT	ACA	CGC	AGC	CGC	CAG	GTG	ACT	TCA	GCC	CAG	CTG	TGG	TTC	CAC	ACC	GGG	CTG	GAC	
Human		Leu	Phe	Arg	Tyr	Met	Phe	Arg	Pro	Ser	Gln	His	Thr	Arg	Ser	Arg	Gln	Val	Thr	Ser	Ala	Gln	Leu	Trp	Phe	His	Thr	Gly	Leu	Asp	
Bovine																															
Porcine																															
Human	432	AGG	CAG	GGC	ACA	GCA	GCC	TCC	AAT	AGC	TCT	GAG	CCC	CTG	CTA	GGC	CTG	CTG	GCA	CTG	TCA	CCG	GGA	CCC	GTG	GCT	GTG	CCC	ATG		
Human		Arg	Gln	Gly	Met	Ala	Ala	TCC	Asn	Ser	Ser	Glu	Pro	Leu	Leu	Gly	Leu	Leu	Val	Ser	Pro	Gly	Gly	Pro	Val	Ala	Val	Pro	Met		
Bovine																															
Porcine																															
Human	519	TCT	TTG	GGC	CAT	GCT	CCC	CCT	CAC	TGG	GCC	GTG	CTG	CAC	CTG	GCC	ACC	TCT	GCT	CTC	TCT	CTG	CTG	ACC	CAC	CCC	GTC	CTG	GTG	CTG	
Human		Ser	Leu	Gly	His	Ala	Pro	Pro	His	Trp	Ala	Val	Leu	His	Leu	Ala	Thr	Ser	Ala	Leu	Ser	Pro	Leu	Leu	Thr	His	Pro	Val	Leu	Val	Leu
Bovine																															
Porcine																															
Human	606	CTG	CTG	CGC	TGT	CCC	CTC	TGT	ACC	TGC	TCA	GCC	CGG	CCT	GAG	GCC	ACG	CCC	TTC	CTG	GTG	GCC	CAC	ACT	CGG	ACC	AGA	CCA	CCC	AGT	
Human		Leu	Leu	Arg	Cys	Pro	Leu	Cys	Thr	Cys	Ser	Ala	Arg	Pro	Glu	Ala	Thr	Pro	Phe	Leu	Val	Ala	His	Thr	Arg	Thr	Ala	Pro	Pro	Ser	
Bovine																															
Porcine																															
Human	693	GGA	GGG	GAG	AGA	GCC	CGA	CGC	TCA	ACT	CCC	CTG	ATG	TCC	TGG	CCT	TGG	TCT	CCC	TCT	GCT	CTG	CGC	CTG	CTG	CAG	AGG	CCT	CCG	GAG	
Human		Gly	Gly	Glu	Arg	Ala	Arg	Arg	Ser	Thr	Pro	Leu	Met	Ser	Trp	Pro	Trp	Ser	Pro	Ser	Ala	Leu	Arg	Leu	Leu	Gln	Arg	Pro	Pro	Glu	
Bovine																															
Porcine																															
Human	780	GAA	CCG	GCT	GCC	CAT	GCC	AAC	TGC	CAC	AGA	GTA	GCA	CTG	AAC	ATC	TCC	TTC	CAG	GAG	CTG	GGC	TGG	GAA	CCG	TGG	ATC	GTG	TAC	CCT	
Human		Glu	Pro	Ala	Ala	His	Ala	Asn	Cys	His	Arg	Val	Ala	Leu	Asn	Ile	Ser	Phe	Gln	Glu	Leu	Gly	Trp	GAA	Arg	Trp	Ile	Val	Thr	Pro	
Bovine																															
Porcine																															
Human	867	CCC	AGT	TTC	ATC	TTC	CAC	TAC	TGT	CAT	GGT	GGT	TGT	GGG	CTG	CAC	ATC	CCA	CCA	AAC	CTG	TCC	CTT	CCA	GTC	CCT	GGG	GCT	CCC	CCT	
Human		Pro	Ser	Phe	Ile	Phe	His	Tyr	Cys	His	Gly	Gly	Cys	Gly	Leu	His	Ile	Pro	Pro	Gln	Asn	Leu	Ser	Pro	Pro	Val	Pro	Ala	Pro	Pro	
Bovine																															
Porcine																															
Human	954	ACC	CCA	GCC	CAG	CCC	TAC	TCC	TTG	CTG	CCA	GGG	GCC	CAG	CCC	TGC	TGT	GCT	GCT	CTC	CCA	GGG	ACC	ATG	AGG	CCC	CTA	CAT	GTC	CGC	
Human		Thr	Pro	Ala	Gln	Pro	Leu	Ser	Leu	Leu	Pro	Gly	Ala	Gln	Pro	Cys	Cys	Ala	Ala	Leu	Pro	Gly	Thr	Met	Arg	Pro	Leu	His	Val	Arg	
Bovine																															
Porcine																															
Human	1041	ACC	ACC	TCG	GAT	GGA	GGT	TAC	TCT	TTC	AAG	TAT	GAG	ACA	GTG	CCC	AAC	CTT	CTC	ACG	CAG	CAC	TGT	GCT	TGT	ATC	TAA	ggg	tgggggg		
Human		Thr	Thr	Ser	Asp	Gly	Gly	Tyr	Ser	Phe	Lys	Tyr	Glu	Thr	Val	Pro	Asn	Leu	Leu	Thr	Gln	His	Cys	Ala	Cys	Ile	***				
Bovine																															
Porcine																															

Fig.1. The nucleotide and encoded amino acid sequences of the human preproinhibin A gene. For comparison, where the bovine and porcine amino acid sequences differ from that of human inhibin the amino acid is shown. (*) Putative *N*-glycosylation sites. (–) Deleted amino acids.

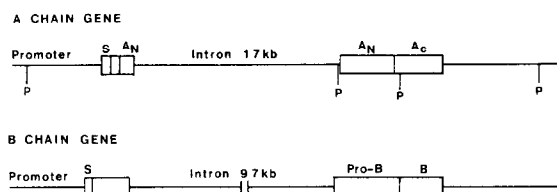


Fig.2. Structures of the human A and B inhibin genes. The regions coding for the preproinhibin chains are boxed. P, *Pst*I restriction sites of the A chain gene. S, putative signal sequences.

4. DISCUSSION

Protein purification studies and DNA cloning techniques have now established the protein structure of bovine, porcine and human inhibins. Mature inhibin is most likely a protein of two disulphide-linked chains A_C and B of molecular masses 20 kDa and 15 kDa, respectively, whereas precursors of inhibin of various molecular masses, notably 58 kDa, have been detected in follicular fluids [3,4]. The cloning of the cDNA genes coding

[illegible]

Fig.3. The nucleotide and encoded amino acid sequences of human inhibin B gene. For comparison, where the nucleotide sequence of the bovine or porcine cDNAs differs the base is shown.

for each of these species of inhibin shows that the two subunits arise from separate genes, as well as identifying the possible precursors of each subunit and the proteolytic cleavage sites [9–11].

Our studies confirm and extend these observations by reporting the first characterisation of human inhibin gene structure at the genomic level. From the Southern blot data the genes coding for the A and B subunits appear to be present as single copies and are not adjacent.

The β_B gene of human inhibin [10] will be described in a later publication. Its low homology with the bovine inhibin B would not permit its identification under the conditions of hybridisation used here.

A striking feature of the similarities between human, bovine and porcine inhibins is that the amino acid sequence of all three B subunits is identical: this represents an extraordinary degree of conservation. The B chains also show similarities to the A chains as well as to Mullerian inhibiting substance, a gonadal protein of importance in sexual differentiation [18], and human transforming growth factor β (TGF β) [19] suggesting a common ancestral gene.

The isolation of genes coding for animal and human inhibins has helped in our understanding of the inhibin structure. The expression of these genes in a suitable host should provide quantities of the protein hitherto unobtainable from the natural source except in trace amounts. This inhibin will be used to give a greater understanding of its physiological role in vivo, to explore the practical applications of inhibin in the manipulation of fertility and to develop sensitive diagnostic systems to evaluate the role of inhibin in human reproductive disorders.

ACKNOWLEDGEMENTS

We wish to thank Professor D.M. de Kretser and Dr D.M. Robertson of Monash University, Professor H. Burger of Prince Henry's Hospital, Dr P. Tolstoshev and Professor N. Willetts for helpful discussions. The work was funded in part by a grant from the National Biotechnology Research Programme of Australia.

REFERENCES

- [1] Baker, H.W.G., Bremner, W.J., Burger, H.G., De Kretser, D.M., Dulmanis, A., Eddie, L.W., Hudson, B., Keogh, E.J., Lee, V.W.K. and Rennie, G.C. (1976) *Recent Prog. Horm. Res.*, **32**, 429–476.
- [2] Franchimont, P., Verstraelen-Proyard, J., Hazee-Hagelstein, M.T., Renard, C.H., Demoulin, A., Bourguignon, J.P. and Husten, J. (1979) *Vitam. Horm.* **37**, 243–302.
- [3] Robertson, D.M., Foulds, L.M., Leversha, L., Morgan, F.J., Hearn, M.T.W., Burger, H.G., Wettenhall, R.E.H. and De Kretser, D.M. (1985) *Biochem. Biophys. Res. Commun.* **126**, 220–226.
- [4] Miyamoto, K., Hasegawa, Y., Fukuda, M., Nomura, M., Igarashi, M., Kangawa, K. and Matsuo, H. (1985) *Biochem. Biophys. Res. Commun.* **129**, 396–403.
- [5] Ling, N., Ying, S.-Y., Ueno, N., Esch, F., Denoroy, L. and Guillemin, R. (1985) *Proc. Natl. Acad. Sci. USA* **82**, 7217–7221.
- [6] Fukuda, M., Miyamoto, K., Hasegawa, Y., Nomura, M., Igarashi, M., Kangawa, K. and Matsuo, H. (1986) *Mol. Cell. Endocrinol.* **44**, 55–60.
- [7] Robertson, D.M., De Vos, F.L., Foulds, L.M., McLachlan, R.I., Burger, H.G., Morgan, F.J., Hearn, M.T.W. and De Kretser, D.M. (1986) *Mol. Cell Endocrinol.* **44**, 271–277.
- [8] Rivier, J., Spiess, J., McClintock, R., Vaughan, J. and Vale, W. (1985) *Biochem. Biophys. Res. Commun.* **133**, 120–127.
- [9] Mason, A.J., Hayflick, J.S., Ling, N., Esch, F., Ueno, N., Ying, S.-Y., Guillemin, R., Niall, H. and Seeburg, P.H. (1985) *Nature* **318**, 659–663.
- [10] Mason, A.J., Niall, H.D. and Seeburg, P.H. (1986) *Biochem. Biophys. Res. Commun.* **135**, 957–964.
- [11] Forage, R.G., Ring, J.M., Brown, R.W., McInerney, B.V., Cobon, G.S., Gregson, R.P., Robertson, D.M., Morgan, F.J., Hearn, M.T.W., Findlay, J.K., Wettenhall, R.E.H., Burger, H.G. and De Kretser, D.M. (1986) *Proc. Natl. Acad. Sci. USA* **86**, 3091–3095.
- [12] Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) in: *Molecular Cloning: A Laboratory Manual*, pp.76–389, Cold Spring Harbor, NY.
- [13] Feinberg, A.P. and Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
- [14] Loenen, W.A.M. and Brammar, W.J. (1980) *Gene* **20**, 249–259.
- [15] Frischauf, A.-M., Lehrbach, H., Poustka, A. and Murray, N. (1983) *J. Mol. Biol.* **170**, 827–842.
- [16] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.

- [17] Mount, S.M. (1982) *Nucleic Acids Res.* 10, 459–472.
- [18] Cate, R.L., Mattaliano, R.J., Hession, C., Tizard, R., Farber, N.M., Cheung, A., Ninfa, E.G., Frey, A.Z., Gash, D.J., Chow, E.P., Fisher, R.A., Bertonis, J.M., Torres, G., Wallner, B.P., Ramachandran, K.L., Ragin, R.C., Manganaro, T.F., MacLaughlin, D.T. and Donahoe, P.K. (1986) *Cell* 45, 685–698.
- [19] Derynck, R., Jarrett, J.A., Chen, E.Y., Eaton, D.H., Bell, J.R., Assoian, R.K., Roberts, A.B., Spann, M.B. and Goeddel, D.V. (1985) *Nature* 316, 701–705.